

FILE COPY

FOREST INSECT LABORATORY  
UNIVERSITY OF CALIFORNIA,  
BERKELEY, CALIFORNIA.

458.1

UNITED STATES DEPARTMENT OF AGRICULTURE

BUREAU OF ENTOMOLOGY AND PLANT QUARANTINE

Project h-1-1

Date

Author

TITLE

MULTIPLE CORRELATION ANALYSES OF  
HAZARD FACTORS ASSOCIATED WITH  
FUTURE INSECT LOSS

By

F. P. Keen  
Senior Entomologist

and

R. C. Grant  
Field Aide

Forest Insect Laboratory  
445 U. S. Court House  
Portland, Oregon  
August 26, 1941

SUBJECT--

INDEX NO.--



Project h-1-1

MULTIPLE CORRELATION ANALYSES OF  
HAZARD FACTORS ASSOCIATED WITH  
FUTURE INSECT LOSS

By

F. P. Keen  
Senior Entomologist

and

R. C. Grant  
Field Aide

Forest Insect Laboratory  
445 U. S. Court House  
Portland, Oregon  
August 26, 1941

## CONTENTS

	Page
INTRODUCTION	1
ANALYSIS I	
(Explanation)	3
Summary of Analysis I	6
ANALYSIS II	
(Explanation)	7
Summary of Analysis II	10
ANALYSIS III	
(Explanation)	11
Summary of Analysis III	14
PART IV. PARTIAL CORRELATION ANALYSIS OF HAZARD FACTORS	15
DISCUSSION	17
REFERENCE REPORTS	18



The problem of determining what forest conditions constitute or add to the hazard of bark beetle damage in the ponderosa pine forests of the Pacific slope and what weight or importance should be given to various hazard factors has engaged the attention of forest entomologists at the Berkeley and Portland laboratories for the past several years.

The recent beetle hazard inventory of ponderosa pine stands in northeastern California conducted by the Berkeley laboratory<sup>3/</sup> has been based upon the assumption that two factors, (1) past loss as indicated by snag counts and (2) the percentage of risk trees in the stand, gave the clue as to what losses from insect attack could be expected in the near future in the five hazard zones established. These two factors of past loss and risk were given equal weight in arriving at an estimate of future loss. Five hazard zones were set up on a more or less empirical basis using knowledge gained from general observations and studies started in 1936.<sup>1/</sup>

A similar attempt to determine area hazard on an empirical basis was discussed by the writer in a report issued July 29, 1939.<sup>2/</sup> Ten factors were recognized as possibly having some influence on hazard and a method was proposed whereby five of these might be combined on an equal basis by a system of transparent map overlays to aid in arriving at an estimate of future hazard. The final answer was a combination of stand data and loss data, interpreted by an entomologist familiar with the forested area in question, and modified to some degree by his personal judgment.

On this basis hazard maps were prepared covering nearly 3,000,000 acres of ponderosa pine in the Klamath and Deschutes Subregions,<sup>4/</sup> and additional hazard maps are being prepared covering other portions of the region.

While all of this work on hazard zonation has been of immediate practical value, still it is empirical in character and largely based on the personal judgment and interpretation of the men doing the zoning job. The question of what factors actually are indicative of future loss and how much weight can be attached to each and in combination with other factors is still unanswered.



The first multiple correlation analysis<sup>2/</sup> was made by the writers using data for 10 years of loss on 25 full section sample plots in the Klamath Basin. This showed that the gross insect loss which occurred on these plots during the 10-year period 1921-1930 was correlated with site index, volume per acre, growth rate and reduction in growth during the period 1921-1930 from the previous average. A regression equation was developed which combined these factors, and with this 51% of the variance in current loss was explained. Of these factors growth reduction was the most important and the only one to show significance in the partial correlation tests. This study, while useful in showing the relation of these factors to current loss, did not indicate how useful they would be in predicting future loss.

If at a given time a survey is made of a certain forest tract and various data taken on site, volume per acre, stand structure or past losses, percentage of risk trees, current growth rate, etc., how closely can one predict the hazard of insect loss occurring in the immediate future? What factor or factors would be useful for such a prediction, which are the more important, and what weight can be given to each?

In looking over the available data taken on the annual regional bark beetle survey throughout eastern Oregon and Washington, we found two sets of data that might be used for such an analysis of future hazard. We had for instance 10 years of records of past loss on 25 640-acre sample plots in the Klamath Basin from 1921 to 1930 inclusive; a stand inventory made in 1932 in which volume per acre, percent of trees in susceptible tree classes, and other data were taken; and then the subsequent loss history for another 10 years on most of these plots. Having all of the past loss history and stand inventory data in 1931, how accurately could the future hazard (i.e., the 1931-1940 losses) be predicted?

Another set of data available for a similar study was loss records taken on 114 320-acre plots scattered throughout eastern Oregon and Washington on which stand inventory data and a snag count had been made in 1937. At that time, could we have predicted the future loss for the next three or four years with a reasonable degree of accuracy? The present analyses were undertaken in the hope of answering some of these questions. Trying various combinations of variables and data four separate multiple correlation analyses were completed. These will be discussed separately and in chronological order.



# ANALYSIS I FUTURE LOSS RELATED TO PAST LOSS, SUSCEPTIBLE VOLUME, AND GROWTH REDUCTION

## 25 Klamath Basin Plots

In a previous study<sup>2/</sup> a multiple correlation analysis was made of the relation of average annual loss in percent of stand killed to site index, volume per acre, normal growth rate in board feet per acre, and percent radial growth reduction during the period 1921-30. Twenty-five full-section plots in the Klamath Basin were used for this study and the results showed that growth reduction was the one outstanding factor correlated with current loss. The other factors all showed high simple correlation but only because of their joint association with other factors directly tied up with loss. As partial correlations their influence on loss was not significant.

The coefficient of multiple correlation in this study was  $R = .8953$  and the standard error of estimate was  $SE_E = .61\%$ .

Using these same data plus the record of loss for the following 10-year period, 1931-1940, an analysis was made to determine how well future loss was correlated with past loss and these other factors.

The following variables were considered:

Dependent:

$(100)x$  = future loss in percent of stand. (Mean annual loss for the 10-year period 1931-1940)

Independent:

$(100)A$  = past mean annual loss in percent of stand for 10-year period 1921-1930

$B$  = susceptible volume per acre in 100 B. Ft. as of 1932. (Includes Dunning's tree classes 4, 5, and 7)

$C$  = radial growth reduction percent. Based on ring width of index trees for last 10-year period (1921-1930) as compared to a mean annual growth of 19.0 mm. for previous decades.



The simple correlation coefficients ( $r$ ) are recorded in the following tabulation:

$x$ - (future loss correlates with)	$r$	Significance
A - (past loss)	+ .6273	High
B - (susceptible vol./acre)	- .5486	High
C - (growth reduction %)	+ .3654	No
A - (past loss correlates with)		
B - (susceptible volume)	- .7949	High
C - (growth reduction %)	+ .6444	High
B - (susceptible volume correlates with)		
C - (growth reduction %)	- .3180	No

There are several interesting deductions to be drawn from an examination of the simple correlation coefficients. We find that future loss is correlated with past loss to a considerable degree. In other words, plots which showed a low loss in 1921-30 continued to show low losses during 1931-40, and high losses continued high except where depletion of stand reversed the trend.

The relation of future loss to susceptible volume per acre showed a significant negative correlation of -.5486. In other words, losses did not occur where the greatest volume of susceptible material occurred, because high volume of susceptible trees was associated with high stand volume, which in turn reflected good site. Poor sites had low stand volume and hence low volume per acre of susceptible material.

The most surprising outcome of the above correlation coefficient determinations was the fact that growth reduction (during 1921-30) showed no significant correlation with future loss (during 1931-40), when in the previous analysis it was the outstandingly significant variable ( $r = +.6444$ ) with current loss. The reason for this apparently was that there was no relation between growth reduction in 1921-30 and 1931-40. Plots which had a severe growth reduction in the first decade were apt to show improvement in the second decade, and low losses, while plots which had shown little reduction in 1921-30 began to show the effect of crowding and stand stagnation in 1931-40 and suffered a proportionately greater loss than they had previously. In other words it was impossible to tell what future loss was going to be, by measuring past growth reduction.



The correlation between independent variables showed a high coefficient of correlation between past loss and susceptible volume (negative) and growth reduction (positive).

Combining the three independent variables of this analysis, we find that the multiple correlation coefficient was  $R = .63$  and was of medium significance (since .552 least significant value; .641 highly significant value) which shows that approximately 40 percent ( $R^2$ ) of the variability in future loss was correlated with these three factors.

The standard error of estimate or the unexplained source of variation in this study is shown in the following table:

Source of Variation	Degrees of Freedom	Sum Squares	Mean Square (Variance)	S.D.
Due to regression	3	8.5033		
Not accounted for	21	13.2665	.6317	.79%
Total	24	21.7698	.9071	.95%

Comparing these results with those of the previous analyses we find:

	1 Correlation of Past Loss with 4 Variables	2 Correlation of Future Loss with 3 Variables
Multiple correlation coefficient	.8953	.63
Percent of variability accounted for	80%	40%
Standard error of estimate (in percent of stand volume killed)	.61%	.79%
Standard deviation reduced by	51%	17%



From the above it is obvious that the three variables used in this problem, having failed to account for 60 percent of the total variation, do not give a very satisfactory measure upon which to base an estimate of future loss.

By comparing the standard regression or "beta" coefficients it is possible to determine the weight or importance of these independent variables in their relation to future loss and whether any one of them is sufficiently useful to be of value. These are shown as follows:

<u>Variable</u>	<u>Standard Regression or Beta Coefficients</u>	<u>Weight in Percentage</u>
Past loss	(A) .5468	79
Susceptible Volume per acre	(B) .1222	18
Growth reduction %	(C) <u>.0257</u>	<u>3</u>
	.6747	100

Of the explained variation in future insect loss, the variable "past loss" accounts for 79% and is by far the most important. Susceptible volume only accounts for 18 percent, while growth reduction has so little influence as to be negligible.

#### Summary of Analysis I

A multiple correlation analysis of the relation of three variables, (1) past loss, (2) susceptible volume and (3) past growth reduction, to future loss showed the following:

1. Significant simple correlations were found between all factors except:
  - (1) Growth reduction and future loss
  - (2) Growth reduction and susceptible volume
2. Multiple correlation coefficient (R) of .63 was of medium and not high significance, and accounts for only 40% of the variability. Not as strong as the R of .89 found in the original analysis.
3. Standard error of estimate ( $SE_E$ ) of .79% out of total standard deviation of .95%. Standard deviation reduced by 17% due to regression. (Mean loss = 2.2672 percent of stand.)
4. Past loss produced 79% of the explained variation in future insect loss in this problem. Susceptible volume 18%, and growth reduction 3%, which was insignificant.



ANALYSIS II  
FUTURE INSECT LOSS RELATED TO PAST LOSS, SUSCEPTIBLE VOLUME, AND  
GROWTH REDUCTION

114 Plots throughout Eastern Oregon and Washington

Since Analysis I gave so little encouragement towards predicting future loss from the three variables tested, another analysis was tried using all regional survey plots throughout eastern Oregon and Washington having a stand inventory taken in 1937 and a continuous loss record for 3 years following this inventory. There were 114 plots found to have comparable records of this character and suitable for use in a multiple correlation analysis.

On these survey plots, mostly 320 acres in size, the past loss was determined from a snag count taken in 1937. The green stand inventory cruise of 1937 showed volume of ponderosa pine per acre and its distribution by Keen's tree classes. Subsequent insect surveys on these plots gave data as to annual loss in percent of stand for the three-year period 1937-39. Growth reduction percents were derived from tree-ring measurement of increment cores taken from "index trees" (Class 3A and 2A) in 1937.

Assuming that we wished to make a prediction in 1937, from the data available at that time, as to what the next three years of insect loss would be, what factors would have been useful and how accurate a prediction would have been possible?

In setting up this problem, scatter diagrams were first prepared. The dependent variable, annual future loss for 1937-39 in percent of stand volume killed was plotted over past loss, susceptible volume per acre in board feet, and in percent of stand, growth reduction for last decade 1927-36 from mean annual growth, and growth reduction for the year 1936. The scatter diagrams indicated the best relationship with past loss, susceptible volume in percent of stand, and growth reduction for 1936 and so a multiple correlation analysis using these factors was set up.



### Dependent Variable:

- (10)  $x$  = Average annual insect loss in percent of stand volume killed during period 1937-1939, as determined from annual 100% insect loss cruises on 114 sample plots.

### Independent Variables:

- A - Past average annual loss in percent of stand killed for approximately a 10-year period (1927-1936), as determined from a 10% snag count and green stand tally in 1937.
- B - Susceptible volume in percent of stand as of 1937 (includes Keen's Classes C, D, and 4B) as determined from a 10% green stand cruise of all plots.
- C - Growth reduction in 1936, percent reduced from mean ring width for period 1850-1937, as determined from tree-ring measurements from 12 "index trees" on each sample plot taken in 1937.

Using the above four variables, the simple correlation coefficients were first computed. These are listed in the following tabulation:

$x$ - (future loss correlates with)	<u><math>r</math></u>	<u>Significance*</u>
A - (past loss)	+ .5688	High
B - (susceptible volume %)	- .1645	None
C - (growth reduction %)	+ .4049	High
A - (past loss correlates with)		
B - (susceptible volume %)	- .0124	None
C - (growth reduction %)	+ .3358	High
B - (susceptible volume % correlates with)		
C - (growth reduction %)	- .3902	High

\*Significant values of  $r$  = .195 least, .254 high



This analysis again shows that past loss is highly correlated with future loss, and growth reduction is again back in the picture after showing no significance in the last analysis. Percent of susceptible volume is not significantly associated with either past or future insect loss, but is negatively correlated with growth reduction percent. Perhaps there is an interrelation between percent of susceptible volume and site or stand competition which would explain this correlation.

The multiple correlation coefficient (R) was computed to be .617 and found to be highly significant with the greater portion of the relationship explained by past loss and growth reduction percent.

<u>Source of Variation</u>	<u>Degrees of Freedom</u>	<u>Sum of Squares</u>	<u>Mean Square Variance</u>	<u>Standard Deviation</u>
Due to regression	3	13.6050	4.5350	
Not accounted for	<u>110</u>	<u>22.0849</u>	<u>.2008</u>	<u>.45%</u>
Total	113	35.6899	.3158	.56%

The regression equation:

$10\bar{x} = .2346A - .0243B + .0842C + 1.87$  was not used to estimate losses but the standard error of estimate was .45% as computed in the above table. This error of estimate was somewhat lower than those computed in the previous analyses because the total range of losses experienced during 1937-1939 was not as large as that found between 1931 and 1940. For instance on these 114 plots the total range in loss was from 0.1 to 2.7 percent of the stand per year--with an average of 0.85%; and S.D.  $\pm .56\%$ .

As the above table indicates, the total standard deviation was 0.56% and through regression this was reduced to 0.45% or a reduction of 20 percent. The percent of variation accounted for was ( $R^2$ ), 38 percent, which still leaves much to be desired.

The relative importance of each variable in explaining future insect loss is brought out in the following table:

<u>Variable</u>	<u>Standard Regression or "Beta" Coefficients</u>	<u>Weight Percentage</u>
Past Loss (A)	.5002	64
Susceptible volume % (B)	.0775	10
Growth reduction % (C)	<u>.2074</u>	<u>26</u>
	.7851	100



As in the previous case, past loss has the greatest weight in explaining future loss, with growth reduction of about  $1/4$  as much significance. Susceptible volume percent was shown to be insignificant.

#### Summary of Analysis II

From a multiple correlation analysis of data taken on 114 320-acre sample plots throughout eastern Oregon and Washington, in which insect loss during the three-year period 1937-39 was related to past loss, percent of susceptible volume in the stand, and the growth reduction shown by the 1936 ring, the following results were obtained:

1. Only past loss and growth reduction showed significant simple correlation coefficients.
2. The multiple correlation coefficient of the three factors with future loss was  $R = .617$ , which was found to be highly significant; and accounted for 38 percent of the variability.
3. Out of the total standard deviation error of .56%, the three factors tested failed to account for .45%. Thus through regression, the standard deviation was reduced by 20 percent.
4. Past loss as shown by snag counts accounted for 64 percent of the explained variation.



ANALYSIS III  
FUTURE INSECT LOSS RELATED TO PAST LOSS, VOLUME PER ACRE,  
AND GROWTH REDUCTION

114 Plots throughout Eastern Oregon and Washington

This study was essentially the same as the preceding except that an attempt was made to substitute some factor of more value than susceptible volume percent. Volume of stand per acre in 100 board feet was chosen because it is one of the best expressions of site which could be used in a problem of this nature.

The same 114 check plots were represented and the dependent variable,  $x$  - average annual insect loss for 3-year period 1937-39, remained the same as in Analysis II. The independent variables included:

- A - (Same as in Analysis II) Past average annual loss in percent of stand killed for approximately a 10-year period (1927-1936), as determined from a 10 percent snag count and green stand tally in 1937.
- C - (Same as in Analysis II) Growth reduction in 1936, percent reduced from mean ring width for period 1850-1937, as determined from tree-ring measurements from 12 "index trees" on each sample plot taken in 1937.
- D - Volume per acre in 100 board feet as determined from a 10% green stand cruise in 1937.

The simple correlation coefficients were first calculated, and it was found that the three independent variables were all significantly correlated with future loss. These were as follows:

$x$ - (future loss correlates with)	<u><math>r</math></u>	<u>Significance*</u>
A (past loss)	+.5688	High
C (growth reduction %)	+.4047	High
D (Volume per acre)	-.3575	High
A - (past loss correlates with)		
C (growth reduction %)	+.3355	High
D (volume per acre)	-.2384	Medium
C - (growth reduction % correlates with)		
D (volume per acre)	-.2394	Medium

\*.195 least significant value of  $r$   
.254 high significant value of  $r$



All of the independent variables correlate more highly with the dependent variable than they do with each other.

The coefficient of multiple correlation was computed and found to be  $R = .642$ , which is highly significant and accounted for 41 percent of the variation. This coefficient shows some slight improvement over  $R = .617$  of the previous analysis.

The multiple regression equation for estimating future loss is as follows:

$$10\bar{X} = .2120A + .0832C - .2539D + 4.1416$$

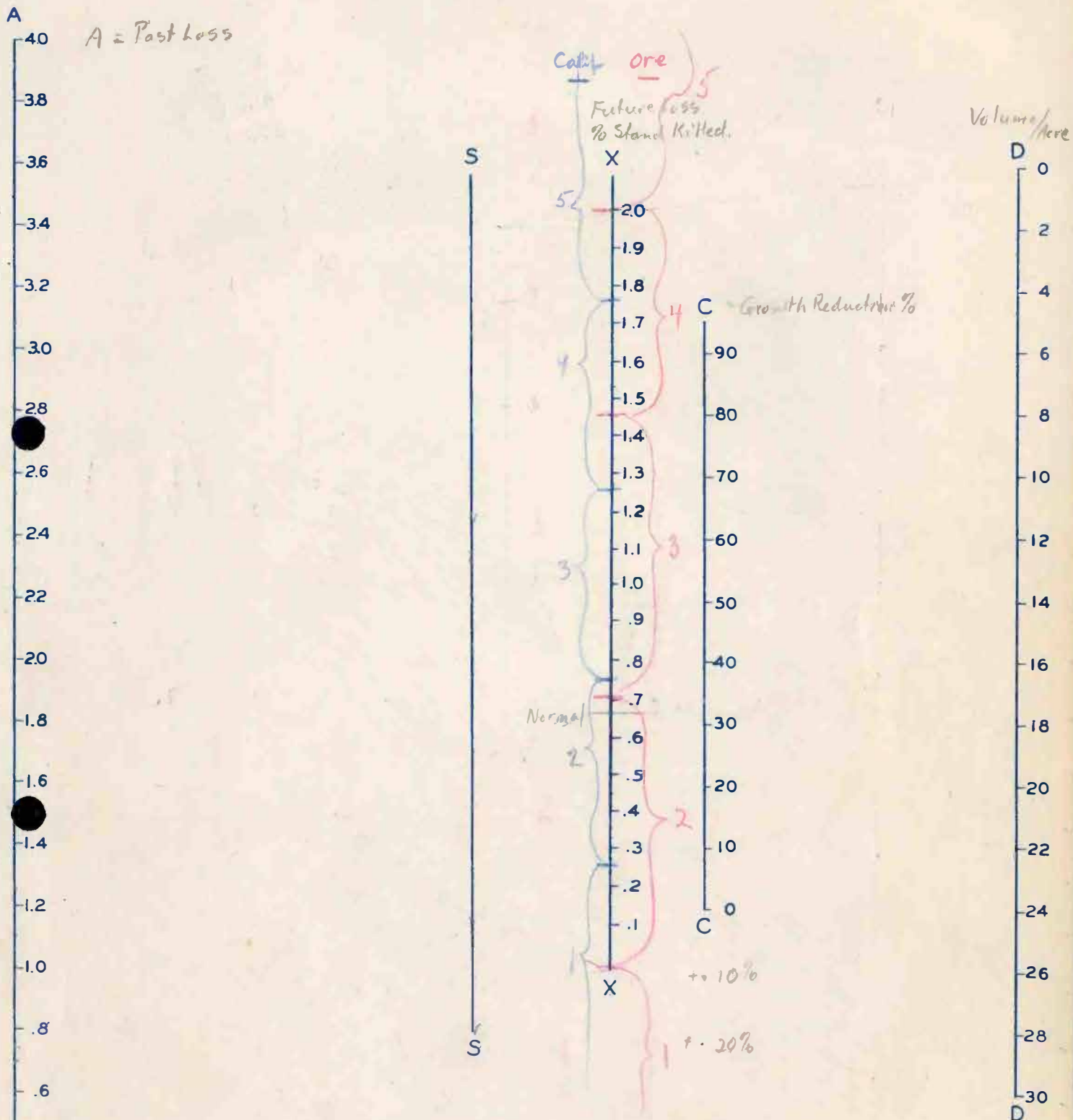
This equation was used to estimate the future loss (1937-39) for the 114 sample plots. The equation was used graphically in the form of an alinement chart (see chart attached) in order to save time in computing the estimates. Each of the three regressions was tested for curvilinearity but none appeared and no improvement in the estimates resulted. The regression lines for each variable were drawn and the estimate residuals plotted about the line. Some refitting of the growth reduction percent and volume per acre curves was attempted and the C and D axes on the alinement chart adjusted from which second estimates were read, but to no avail. The x axis was regraduated by plotting the measured values over the first estimates and reading the new values from the curve. The third estimates were then read from the alinement chart but these were no better, so the chart was put back in its original form. The SEg was .44% stand killed, which is some improvement over previous estimates.

An analysis of the variation in the multiple regression is given in the following table:

<u>Source of Variation</u>	<u>Degrees Freedom</u>	<u>Sum of Squares</u>	<u>Mean Square Variance</u>	<u>S.D.</u>
Due to regression	3	14.7007	4.9002	
Not accounted for	110	20.9893	.1908	.44%
Total	113	35.6900	.3158	.56%



# ALINEMENT CHART FOR HAZARD FACTORS (PAST LOSS, VOL./ACRE, GROWTH REDUCTION)



FOR EQUATION:  $10\bar{x} = .2120 A + .0832 C - .2539 D + 4.1416$

RULE: FROM A TO C ; HOLD S ; TO D ; READ X

A = PAST LOSS - % OF STAND KILLED ANNUALLY

C = GROWTH REDUCTION % (1936)

D = VOLUME / ACRE

X = FUTURE LOSS - % OF STAND KILLED ANNUALLY



The standard error of estimate ( $SE_E$ ) compares with that computed from the alignment chart by taking the root mean square of the deviations about the line of estimate. The substitution of volume per acre for susceptible volume has not explained much more of the variation in insect loss. Simple correlation coefficients in the early stages of this problem indicated that improvement in the estimate might be expected, but evidently the independent variables are so interrelated and associated in their effect on loss that the addition or deletion of any one variable does little to change the estimate one way or another. Although in this last analysis we have less variation in loss not associated with the variables we still have explained only 41% of the variation.

The relative importance and the significance of the three variables are shown in the following table:

<u>Variable</u>	<u>Standard Regression or Beta Coefficient</u>	<u>Weight in Percent</u>	<u>S.D. Betas</u>	<u>"t"</u>
Past loss (A)	.4522	53	.0787	5.7459
Growth reduction (C)	.2050	24	.0787	2.6048
Volume per Acre (D)	-.2006	23	.0762	2.6325
	.8578	100		

\*Significant values of "t": low (5%) - 1.984; high (1%) - 2.626

Growth reduction is not highly significant while volume per acre is just barely so. These borderline cases justify a more thorough analysis of the factors as brought out by partial correlation coefficients. Past loss continues to account for the greatest part of the variation in future loss while growth reduction and volume per acre are almost equal in their effects on future loss.



### Summary of Analysis III

A multiple correlation analysis of the hazard factors associated with future loss showed:

1. That past loss, 1936 growth reduction, and volume per acre gave highly significant simple correlation coefficients with the subsequent (future) loss of 1937-39.
2. That these three independent variables gave a multiple correlation coefficient of 0.64, which accounted for 41 percent of the variability.
3. A total standard deviation in future loss of 0.56% of the stand, out of which the standard error of estimate was 0.44%. Thus the standard deviation was reduced by 21 percent due to regression.
4. Two factors, (1) past loss and (2) volume per acre, were highly significant, while growth reduction was merely significant.
5. Of the portion of variability explained by regression, past loss accounted for 53 percent, while growth reduction and volume per acre accounted for 24 percent and 23 percent respectively.



# PART IV PARTIAL CORRELATION ANALYSIS OF HAZARD FACTORS

In the preceding problem it was proven that the three factors of past loss, growth reduction, and volume per acre all exerted some influence on future loss, and explained 41% of the variation in loss. There is the possibility in this analysis that the variables may be associated with one another and that perhaps two or even only one of the three may be significantly responsible when shorn of relationship with other independent variables. Partial correlation is the technique used to determine the relationship between two variables when the effects of their common association with other variables are partially removed.

The first step in this final analysis consisted of computing the first order partial correlation coefficients. These coefficients express the relationship between any two independent variables irrespective of the common effect of the dependent variable.

$$\begin{array}{ll} r_{AC \cdot x} = .1401 & r_{AC} = .3355 \\ r_{AD \cdot x} = -.0457 & r_{AD} = -.2384 \\ r_{DC \cdot x} = -.1107 & r_{DC} = -.2394 \end{array}$$

$$\text{Equation: } r_{AC \cdot x} = \frac{r_{AC} - (r_{Ax})(r_{Cx})}{\sqrt{1-r_{Ax}^2} \cdot \sqrt{1-r_{Cx}^2}}$$

.195 significant

.254 highly significant

A comparison of the first order partial correlation values with the simple correlation values readily shows that the dependent variable is responsible for significant simple correlation between the independent variables. When the common association of the independent variables with future loss is discounted no significant correlation exists between past loss, 1936 growth reduction, and volume per acre.

We now know that x is more highly associated with A, C, and D than they are between themselves. But how much influence does each variable exert on future loss when all other relationships are nullified? The individual effect of each factor can be determined by computing the fourth order partial correlation coefficients. This process is essentially the same procedure as followed previously in solving the normal equations except that each independent variable is substituted as the dependent.

After the Betas are determined the following formula is used to find the partial correlation coefficient:

$$r_{Dx \cdot AC} = \text{square root of } \beta_{x D} \cdot \beta_{D x}$$

The above factor is the correlation between volume per acre and future loss independent of the effects of past loss and growth reduction. The fourth order partials for the relationship between future loss and each independent variable are as follows:



### Partial Coefficients

$$r_{Dx \cdot AC} = .2427$$

$$r_{Cx \cdot AD} = .2407$$

$$r_{Ax \cdot CD} = .4800$$

Significant values are .254 1%  
.195 5%

All of the factors are significantly associated with future loss and one, past loss, is highly significant. Volume per acre and 1936 growth reduction are correlated about equally with future loss, but past loss is more closely associated with future loss, and, in this problem, is responsible for half of the explained variation.

The fourth order partial coefficient for  $r_{DC \cdot Ax} = .1058$ . The first order  $r_{DC \cdot x} = -.1138$  is slightly higher and comparison of the two coefficients indicates that past loss accounts for some of the relationship between DC, but the association is slight.

The partial coefficients bring out the fact that there exists very little intercorrelation between the independent variables, none of which is significant, and that the significance of the simple correlation coefficients is due to the common association between the independent variables and the dependent.



## DISCUSSION

Of the 10 factors previously considered<sup>2/</sup> as possibly having some relation to pine beetle hazard, three have been found to be definitely significant. These three independent variables, (1) past losses, (2) stand volume per acre, and (3) growth reduction, have been combined in a multiple regression equation and alinement chart so that future loss can be estimated on the basis of combining these three factors. The result can be expected to explain at least 41 percent of the variability of future loss.

Stand structure, at least percent of susceptible tree classes in a stand, has not been shown to be significantly correlated with future loss and hence could not be used.

Six of the other factors mentioned<sup>2/</sup> have not been tested. Of these (1) life zone, (2) elevation, (3) temperature, (4) forest type and (5) site quality are all quite similar in being different expressions for life zone or site quality.

Most of these can not be expressed mathematically except in broad classes. Site quality can be expressed as site index to obtain a greater refinement in subdivision of classes. No measurements of temperature are available. Preliminary tests through means of scatter diagrams gave no indication that "past fire history" would be useful in this connection.

To complete the present series of correlation analyses, computations should be made to determine the significance and value of life zone or site index in a multiple regression equation. When this is done the present series of multiple correlation analyses may be considered complete.



## REFERENCE REPORTS

1. Salman, K. A. and Johnson, P. C.  
The Forest Insect Hazard Inventory of Eastside Forest Areas.  
Preliminary report on methods and their application.  
(Unpublished report) Sept. 21, 1937.
2. Keen, F. P.  
Multiple correlation analysis of gross insect loss to hazard  
factors on Klamath Basin plots, Oregon. June 20, 1939.
3. Keen, F. P.  
An approach to the problem of determining pine beetle hazard  
in ponderosa pine forests of Oregon and Washington.  
July 29, 1939.
4. Keen, F. P.  
A preliminary analysis of pine-beetle hazard conditions in  
the Klamath Subregion. July 31, 1939.
5. Miller, J. M., Salman, K. A., and Johnson, P. C.  
Bark beetle hazards in the pine stands of northeastern  
California. May 15, 1941.



# ALINEMENT CHART FOR HAZARD FACTORS ( PAST LOSS, VOL./ACRE, GROWTH REDUCTION )

